INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

Statistics on Income and Living Conditions – EU-SILC
Action 2009

FINAL QUALITY REPORT

Departamento de Estatísticas Demográficas e Sociais
Serviço de Estatísticas das Condições de Vida

December 2011

1. COMMON LONGITUDINAL EUROPEAN UNION INDICATORS BASED ON THE LONGITUDINAL COMPONENT OF EU-SILC

For reasons of clarity, i.e. avoiding any source of misinterpretation, we consider that all point or annual estimates shall be based on the cross-sectional component of EU-SILC, including PY080 (from 2006 to 2009).

2. ACCURACY

2.1 Sample design

Sections 2.1.1 to 2.1.8.3 and 2.1.8.8 to 2.1.9 do not apply.

2.1.8.4 Final longitudinal weight

In order to calculate the base weights it is necessary to introduce the concept of re-entries. These units correspond to sample persons not present in the second year of a three consecutive years period. According to the scheme presented on page 42 of the document EU-SILC 065 (2009 operation) the population can be divided into five sets: A, B, C, D and E.

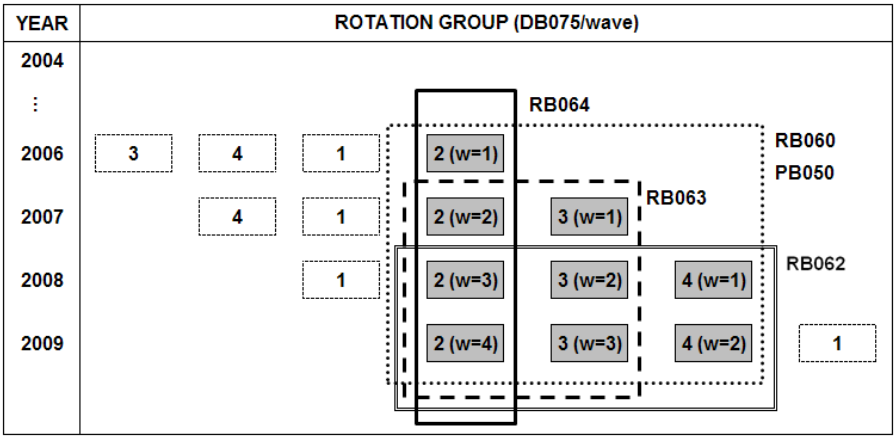
Figure 1 – Re-entries scheme

Population (w=1)	Sample (w=1)	Sample (w=2)	Sample (w=3)
A	a	a	a
B	b	b	
C	c		c
D	d		
E			
sample →	aχbχcχd	aχb	aχc

A, a = part of the population/sample which potentially responds at all three waves, w=1 to 3.;
B, b = potential respondents at w=1 and w=2, but not at w=3;
C, c = potential respondents at w=1 and w=3, but not at w=2 (re-entries);
D, d = potential respondents at w=1, but not at any subsequent wave;
E = potential non-respondents at w=1 (they are not followed).

The longitudinal component involves the calculation of five individual weights: RB060, PB050, RB062, RB063 and RB064.

Figure 2 – Rotation scheme



RB060 and PB050: personal base weight

Three cases have to be distinguished:

- Panel selected in 2006 (DB075=2 → Longitudinal set of four years duration, from 2006 to 2009);
- Panel selected in 2007 (DB075=3 → Longitudinal set of three years duration, from 2007 to 2009);
- Panel selected in 2008 (DB075=4 → Longitudinal set of two years duration, from 2008 to 2009).

Panel selected in 2006 (longitudinal set of four years duration, from 2006 to 2009).

a) Year 2006 (w=1)

The individual weights RB060 (base weights) are equal to the cross-sectional weights RB050, multiplied by 4 to take into account the fact that each rotational group, for each wave, should represent the longitudinal population in scope in 2006.

$$RB060_{i\,(w=1)}^{2006} = 4 \cdot RB050_{i\,(w=1)}^{2006}$$

For persons aged 16 or more the weight PB050 is equal to the cross-sectional weight PB040:

$$PB050_{i\,(w=1)}^{2006} = 4 \cdot PB040_{i\,(w=1)}^{2006}$$

b) Year 2007 (w=2)

The weights RB060 are obtained multiplying the base weights calculated in 2006, inflated taking into account the attrition:

$$RB060_{i\,(w=2)}^{2007} = RB060_{i\,(w=1)}^{2006} \cdot \frac{1}{\hat{p}_{1,h}^{2006\to2007}},$$

where the estimated response probability within region h in the second wave is given by:

$$\hat{p}_{1,h}^{2006\to2007} = \frac{n_{a\cup b,h}^{2007}}{n_{a\cup b\cup c\cup d,h}^{2006} - m_{1,h}^{2006\to2007}},$$

and $n_{a\cup b,h}^{2007}$ and $n_{a\cup b\cup c\cup d,h}^{2006}$ are, respectively, the number of sample persons enumerated in 2007 and in 2006. The quantity $m_{1,h}^{2006\to2007}$ denotes the number of sample persons out-of-scope between 2006 and 2007 in region *h*.

"Out-of-scope" situations correspond to persons who have moved to a collective household, to a foreign country, died or were unable to locate.

Children born to a sample woman receive the weight of the mother; persons moving into sample households from non-sample households in the population (co-residents) or having left the population are given zero weights; former household members receive a zero weight.

The base weight for persons aged 16 or more (PB050) is obtained multiplying the base weight RB060 by a re-weighting factor calculated by region, sex and five year age-groups (the same used in calibration). In each cell (C) the response probability is given as the ratio between the sum of the base weights RB060 of persons who have replied the individual questionnaire (Q) and the sum of the base weights RB060 for all individuals.

$$PB050_{i\,(w=2)}^{2007} = RB060_{i\,(w=2)}^{2007} \cdot \frac{1}{\hat{p}_C^{2007}},$$

where

$$\hat{p}_C^{2007} = \frac{\sum_{i\in c\cap Q} RB060_{i\,(w=2)}^{2007}}{\sum_{i\in C} RB060_{i\,(w=2)}^{2007}}.$$

c) Year 2008 (w=3)

Two cases have to be distinguished: if the sample person belongs to set "a" or set "c" in figure 1.

If the sample person was enumerated in 2006 and 2007 (set "a"), the base weight RB060 is calculated multiplying the base weight of the previous year by a factor that takes into account the attrition between 2007 and 2008 and by a factor to compensate the re-entries.

$$RB060_{i\,(w=3)}^{2008} = RB060_{i\,(w=2)}^{2007} \cdot \frac{1}{\hat{p}_{2,h}^{2007\to2008}} \cdot \frac{n_{a,h}^{2008}}{n_{a,h}^{2008} + n_{c,h}^{2008}},$$

with

$$\hat{p}_{2,h}^{2007\to2008} = \frac{n_{a,h}^{2008}}{n_{a\cup b,h}^{2007} - m_{2,h}^{2007\to2008}},$$

where $n_{a,h}^{2008}$ and $n_{c,h}^{2008}$ are the number of sample persons enumerated in 2008 in sets "a" and "c", respectively. The quantity $n_{a\cup b,h}^{2007}$ denotes the number of sample persons in 2007 and $m_{2,h}^{2007\to2008}$ the number of sample persons out-of-scope between 2007 and 2008 in region *h*.

If the sample person was a respondent in 2006 and 2008, but not in 2007 (re-entries), the weight RB060 is obtained multiplying the base weight calculated in 2006, inflated taking into account the attrition between 2006 and 2008:

$$RB060_{i\,(w=3)}^{2008} = RB060_{i\,(w=1)}^{2006} \cdot \frac{1}{\hat{p}_{3,h}^{2006\to2008}}.$$

with

$$\hat{p}_{3,h}^{2006\to2008} = \frac{n_{a,h}^{2008} + n_{c,h}^{2008}}{n_{a\cup b\cup c\cup d,h}^{2006} - m_{3,h}^{2006\to2008}},$$

where $n_{a\cup b\cup c\cup d,h}^{2006}$ is the number of sample persons in 2006. The quantity $m_{3,h}^{2006\to2008}$ denotes the number of sample persons out-of-scope between 2006 and 2008 in region *h*.

The base weight for persons aged 16 or more (PB050) is calculated as described above in section b).

$$PB050_{i\,(w=3)}^{2008} = RB060_{i\,(w=3)}^{2008} \cdot \frac{1}{\hat{p}_C^{2008}},$$

with,

$$\hat{p}_C^{2008} = \frac{\sum_{i\in c\cap Q} RB060_{i\,(w=3)}^{2008}}{\sum_{i\in C} RB060_{i\,(w=3)}^{2008}},$$

d) Year 2009 (w=4)

It should be noted that units in the fourth wave were present in both first and second waves (due to the following rules). The analysis is analogue to the previous case and again two cases have to be distinguished: if the sample person belongs to set "a" or set "c" as indicated in figure 1.

If the sample person was enumerated in 2008 (set "a") the base weight RB060 is calculated multiplying the base weight of the previous year by a factor that takes into account the attrition between 2008 and 2009 and by a factor to compensate the re-entries.

$$RB060_{i\,(w=4)}^{2009} = RB060_{i\,(w=3)}^{2008} \cdot \frac{1}{\hat{p}_{4,h}^{2008\to2009}} \cdot \frac{n_{a,h}^{2009}}{n_{a,h}^{2009} + n_{c,h}^{2009}},$$

with

$$\hat{p}_{4,h}^{2008\to2009} = \frac{n_{a,h}^{2009}}{n_{a\cup b,h}^{2008} - m_{4,h}^{2008\to2009}},$$

where $n_{a,h}^{2009}$ and $n_{c,h}^{2009}$ are the number of sample persons enumerated in 2009 in set "a" and set "c", respectively. The quantity $n_{a\cup b,h}^{2008}$ denotes the number of sample persons in 2008 and $m_{4,h}^{2008\to2009}$ the number of sample persons out-of-scope between 2008 and 2009 in region *h*.

If the sample person was a non-respondent in 2008 the base weights RB060 are obtained multiplying the base weights calculated in 2007, inflated taking into account the attrition between 2007 and 2009:

$$RB060_{i\,(w=4)}^{2009} = RB060_{i\,(w=2)}^{2007} \cdot \frac{1}{\hat{p}_{5,h}^{2007\to2009}},$$

with

$$\hat{p}_{5,h}^{2007\to2009} = \frac{n_{a,h}^{2009} + n_{c,h}^{2009}}{n_{a\cup b\cup c\cup d,h}^{2007} - m_{5,h}^{2007\to2009}},$$

where $n_{a\cup b\cup c\cup d,h}^{2007}$ is the number of sample persons in 2007. The quantity $m_{3,h}^{2007\to2009}$ denotes the number of sample persons out-of-scope between 2007 and 2009 in region *h*.

The base weight for persons aged 16 or more (PB050) is calculated as described above in section b).

$$PB050_{i\,(w=4)}^{2009} = RB060_{i\,(w=4)}^{2009} \cdot \frac{1}{\hat{p}_C^{2009}},$$

with

$$\hat{p}_C^{2009} = \frac{\sum_{i\in C\cap Q} RB060_{i\,(w=4)}^{2009}}{\sum_{i\in C} RB060_{i\,(w=4)}^{2009}}.$$

Panel selected in 2007 (longitudinal set of three years duration, from 2007 to 2009).

a) Year 2007 (w=1)

The individual weights RB060 (base weights) are equal to the cross-sectional weights RB050, multiplied by 4 to take into account the fact that each rotational group, for each wave, should represent the longitudinal population in scope in 2007.

$$RB060_{i\,(w=1)}^{2007} = 4 \cdot RB050_{i\,(w=1)}^{2007}$$

For persons aged 16 or more the weight PB050 is equal to the cross-sectional weight PB040:

$$PB050_{i\,(w=1)}^{2007} = 4 \cdot PB040_{i\,(w=1)}^{2007}$$

b) Year 2008 (w=2)

The weights RB060 are obtained multiplying the base weights calculated in 2007, inflated taking into account the attrition:

$$RB060_{i\,(w=2)}^{2008} = RB060_{i\,(w=1)}^{2007} \cdot \frac{1}{\hat{p}_{1,h}^{2007\to2008}}.$$

where the estimated response probability within region *h* in the second wave is given by:

$$\hat{p}_{1,h}^{2007\to2008} = \frac{n_{a\cup b,h}^{2008}}{n_{a\cup b\cup c\cup d,h}^{2007} - m_{1,h}^{2007\to2008}},$$

and $n_{a\cup b,h}^{2008}$ and $n_{a\cup b\cup c\cup d,h}^{2007}$ are, respectively, the number of sample persons enumerated in 2008 and in 2007. The quantity $m_{1,h}^{2007\to2008}$ denotes the number of sample persons out-of-scope between 2007 and 2008 in region *h*.

"Out-of-scope" situations correspond to persons who have moved to a collective household, to a foreign country, died or were unable to locate.

Children born to a sample woman receive the weight of the mother; persons moving into sample households from another non-sample households in the population (co-residents) or having left the population are given zero weights; former household members receive a zero weight.

The base weight for persons aged 16 or more (PB050) is obtained multiplying the base weight RB060 by a re-weighting factor calculated by region, sex and five year age-groups (the same used in calibration). In each cell (C) the response probability is given as the ratio between the sum of the base weights RB060 of persons who have replied the individual questionnaire (Q) and the sum of the base weights RB060 for all individuals.

$$PB050_{i\,(w=2)}^{2008} = RB060_{i\,(w=2)}^{2008} \cdot \frac{1}{\hat{p}_C^{2008}},$$

where

$$\hat{p}_C^{2008} = \frac{\sum_{i \in C \cap Q} RB060_{i\,(w=2)}^{2008}}{\sum_{i \in C} RB060_{i\,(w=2)}^{2008}}.$$

c) Year 2009 (w=3)

Two cases have to be distinguished: if the sample person belongs to set "a" or set "c" in figure 1.

If the sample person was enumerated in 2007 and 2008 (set "a"), the base weight RB060 is calculated multiplying the base weight of the previous year by a factor that takes into account the attrition between 2008 and 2009 and by a factor to compensate the re-entries.

$$RB060_{i\,(w=3)}^{2009} = RB060_{i\,(w=2)}^{2008} \cdot \frac{1}{\hat{p}_{2,h}^{2008\to2009}} \cdot \frac{n_{a,h}^{2009}}{n_{a,h}^{2009} + n_{c,h}^{2009}},$$

with

$$\hat{p}_{2,h}^{2008\to2009} = \frac{n_{a,h}^{2009}}{n_{a\cup b,h}^{2008} - m_{2,h}^{2008\to2009}},$$

where $n_{a,h}^{2009}$ and $n_{c,h}^{2009}$ are the number of sample persons enumerated in 2009 in sets "a" and "c", respectively. The quantity $n_{a\cup b,h}^{2008}$ denotes the number of sample persons in 2008 and $m_{2,h}^{2008\to2009}$ the number of sample persons out-of-scope between 2008 and 2009 in region h.

If the sample person was a respondent in 2007 and 2009, but not in 2008 (re-entree), the weight RB060 is obtained multiplying the base weight calculated in 2007, inflated taking into account the attrition between 2007 and 2009:

$$RB060_{i\,(w=3)}^{2009} = RB060_{i\,(w=1)}^{2007} \cdot \frac{1}{\hat{p}_{3,h}^{2007\to2009}},$$

with

$$\hat{p}_{3,h}^{2007\to2009} = \frac{n_{a,h}^{2009} + n_{c,h}^{2009}}{n_{a\cup b\cup c\cup d,h}^{2007} - m_{3,h}^{2007\to2009}},$$

where $n_{a\cup b\cup c\cup d,h}^{2007}$ is the number of sample persons in 2007. The quantity $m_{3,h}^{2007\to2009}$ denotes the number of sample persons out-of-scope between 2007 and 2009 in region h.

The base weight for persons aged 16 or more (PB050) is calculated as described above in section b).

$$PB050_{i\,(w=3)}^{2009} = RB060_{i\,(w=3)}^{2009} \cdot \frac{1}{\hat{p}_C^{2009}},$$

with,

$$\hat{p}_C^{2009} = \frac{\sum_{i \in C \cap Q} RB060_{i\,(w=3)}^{2009}}{\sum_{i \in C} RB060_{i\,(w=3)}^{2009}}.$$

Panel selected in 2008 (longitudinal set of two years duration, from 2008 to 2009).

a) Year 2008 (w=1)

The individual weights RB060 (base weights) are equal to the cross-sectional weights RB050, multiplied by 4 to take into account the fact that each rotational group, for each wave, should represent the longitudinal population in scope in 2008.

$$RB060_{i\,(w=1)}^{2008} = 4 \cdot RB050_{i\,(w=1)}^{2008}$$

For persons aged 16 or more the weight PB050 is equal to the cross-sectional weight PB040:

$$PB050_{i\,(w=1)}^{2008} = 4 \cdot PB040_{i\,(w=1)}^{2008}$$

b) Year 2009 (w=2)

The weights RB060 are obtained multiplying the base weights calculated in 2008, inflated taking into account the attrition:

$$RB060_{i\,(w=2)}^{2009} = RB060_{i\,(w=1)}^{2008} \cdot \frac{1}{\hat{p}_{1,h}^{2008 \to 2009}}.$$

where the estimated response probability within region *h* in the second wave is given by:

$$\hat{p}_{1,h}^{2008 \to 2009} = \frac{n_{a \cup b,h}^{2009}}{n_{a \cup b \cup c \cup d,h}^{2008} - m_{1,h}^{2008 \to 2009}},$$

and $n_{a \cup b,h}^{2009}$ and $n_{a \cup b \cup c \cup d,h}^{2008}$ are, respectively, the number of sample persons enumerated in 2009 and in 2008. The quantity $m_{1,h}^{2008 \to 2009}$ denotes the number of sample persons out-of-scope between 2008 and 2009 in region *h*.

"Out-of-scope" situations correspond to persons who have moved to a collective household, to a foreign country, died or were unable to locate.

Children born to a sample woman receive the weight of the mother; persons moving into sample households from another non-sample households in the population (co-residents) or having left the population are given zero weights; former household members receive a zero weight.

The base weight for persons aged 16 or more (PB050) is obtained multiplying the base weight RB060 by a re-weighting factor calculated by region, sex and five year age-groups (the same used in calibration). In each cell (*C*) the response probability is given as the ratio between the sum of the base weights RB060 of persons who have replied the individual questionnaire (*Q*) and the sum of the base weights RB060 for all individuals.

$$PB050_{i\,(w=2)}^{2009} = RB060_{i\,(w=2)}^{2009} \cdot \frac{1}{\hat{p}_C^{2009}},$$

where

$$\hat{p}_C^{2009} = \frac{\sum_{i \in C \cap Q} RB060_{i\,(w=2)}^{2009}}{\sum_{i \in C} RB060_{i\,(w=2)}^{2009}}.$$

RB062: Longitudinal weight of two-year duration, for the most recent period 2008 to 2009

a) Years 2006, 2007 and 2008

RB062 is missing as these years do not correspond to the last wave.

$$RB062_{i\,(w=3)}^{2008} = missing,$$

$$RB062_{i\,(w=2)}^{2007} = missing,$$

$$RB062_{i\,(w=1)}^{2006} = missing.$$

b) Year 2009

The sum of the weights of the panels selected in 2006, 2007 and 2008 should be the longitudinal population in scope from the first year of the panel till the current year.

Panel selected in 2006 (wave 4)

$$RB062_{i\,(w=4)}^{2009} = \frac{1}{3} \cdot RB060_{i\,(w=4)}^{2009}.$$

Panel selected in 2007 (wave 3)

$$RB062_{i\,(w=3)}^{2009} = \frac{1}{3} \cdot RB060_{i\,(w=3)}^{2009}.$$

Panel selected in 2008 (wave 2)

$$RB062_{i\,(w=2)}^{2009} = \frac{1}{3} \cdot RB060_{i\,(w=2)}^{2009}.$$

Members with RB110 = 3, 4, 5, 6 or 7 (moved into from outside sample, newly born, moved out, died or not in register) have a zero weight.

RB063: Longitudinal weight of three-year duration, for 2007 to 2009

o **Years 2006, 2007 and 2008**

RB063 is missing as these years do not correspond to the last wave.

$$RB063_{i\,(w=3)}^{2008} = missing,$$

$$RB063_{i\,(w=2)}^{2007} = missing\,,$$

$$RB063_{i\,(w=1)}^{2006} = missing\,.$$

- **Year 2009**

The sum of the weights of the panels selected in 2006 and 2007 should be the longitudinal population in scope from the first year of the panel till the current year.

Panel selected in 2006 (wave 4)

$$RB063_{i\,(w=4)}^{2009} = \frac{1}{2} \cdot RB060_{i\,(w=4)}^{2009}.$$

Panel selected in 2007 (wave 3)

$$RB063_{i\,(w=3)}^{2009} = \frac{1}{2} \cdot RB060_{i\,(w=3)}^{2009}.$$

Members with RB110 = 3, 4, 5, 6 or 7 (moved into from outside sample, newly born, moved out, died or not in register) have a zero weight.

RB064: Longitudinal weight of four-year duration, for 2006 to 2009

- **Years 2006, 2007 and 2008**

RB064 is missing as these years do not correspond to the last wave.

$$RB064_{i\,(w=3)}^{2008} = missing\,,$$

$$RB064_{i\,(w=2)}^{2007} = missing\,,$$

$$RB064_{i\,(w=1)}^{2006} = missing\,.$$

- **Year 2009**

The sum of the weights of the panel selected in 2006 should be the longitudinal population in scope from the first year of the panel till the current year.

Panel selected in 2006 (wave 4)

$$RB064_{i\,(w=4)}^{2009} = RB060_{i\,(w=4)}^{2009}.$$

Members with RB110 = 3, 4, 5, 6 or 7 (moved into from outside sample, newly born, moved out, died or not in register) have a zero weight.

2.1.8.7 Final longitudinal weight

This section corresponds to 2.1.8.4 section.

2.2 Sampling errors

No specific imputation was processed for the longitudinal component, i.e., longitudinal data corresponds to the common rotations cross-sectional data plus the cross-sectional data on emigrants, deceased, lost and former members. For that reason, and in view with consistency with the EU-SILC 2009 indicators, information requested concerns the cross-sectional data.

EU-SILC 2009 cross-sectional				
Income components	Mean (weighted)	Number of		Standard error (weighted)
		Before imputation	After imputation	
Total disposable household income (HY010)	22984	(a)	4961	586
Total disposable household income (HY020)	18380	(a)	4961	412
Total disposable household income before social transfers other than old-age and survivors' benefits (HY022)	17592	(a)	4903	420
survivors' benefits (HY023)	15022	(a)	4229	424
(a) Total disposable household income corresponds to the sum of various components, independently of the pattern of gross/net collection and imputation/no imputation. It is a final step using component series reflecting heterogeneous methods of imputation both in terms of algorithms and number of observations. Because of this, all imputation flags associated with HY020, HY022 and HY023 inform about a mixture of net and gross collection values and an imputation factor of 1.				

EU-SILC 2009 cross-sectional				
Income components	Mean (weighted)	Number of observations not null		Standard error (weighted)
		Before imputation	After imputation (b)	
Income from rental of property or land (HY040G)	5261	260	260	481
Family/children-related allowances (HY050G)	765	0	1254	34
Social exclusion payments not elsewhere classified (HY060G)	2960	0	113	272
Housing allowances (HY070G)	485	0	215	21
Regular inter-household cash transfers received (HY080G)	4831	0	160	732
Interest, dividends, profit from capital investment in unincorporated businesses (HY090G)	1715	532	532	331
Income received by people aged under 16 (HY110G)	1179	0	4	267
Regular taxes on wealth (HY120G)	255	0	2438	12
Regular inter-household transfers paid (HY130G)	3040	0	142	398
Note: data on variables HY050, HY060, HY070, HY080, HY110, HY120 and HY130 was collected net, implying that all gross data was not available before imputation.				

EU-SILC 2009 cross-sectional				
Income components	Mean (weighted)	Number of observations not null		Standard error (weighted)
		Before imputation	After imputation (b)	
Cash or near-cash employee income (PY010G)	13149	1123	4653	316
Cash profits or losses from self-employment (PY050G)	13245	935	935	1220
Pension from individual private plans (PY080)	5230	0	44	1818
Unemployment benefits (PY090G)	3946	0	314	181
Old-age benefits (PY100G)	7415	3327	3327	246
Survivors' benefits (PY110G)	3232	779	779	114
Sickness benefits (PY120G)	2912	0	136	271
Disability benefits (PY130G)	4293	358	358	230
Education-related allowances (PY140G)	1602	0	71	154

(b) Imputation includes partial imputation when one or more of the questions associated to the component are missing, conversion of data collected from net to gross, and total imputation of net data when all the questions associated with the component are missing.

Note: data on variables PY090, PY120 and PY140 was collected net, implying that all gross data was not available before imputation.

EU-SILC 2009 cross-sectional					
Income components		Mean (weighted)	Number of observations		Standard error (weighted)
			Before imputation	After imputation	
Subclasses by household type					
	1 household member	9378	(a)	975	352
	2 household member	10680	(a)	3356	339
	3 household member	10922	(a)	3273	297
	4 and more	10013	(a)	5409	439
Population by age group					
	<25	9512	(a)	3231	346
	25 - 34	11627	(a)	1337	573
	35 - 44	10171	(a)	1707	330
	45 - 54	10905	(a)	1932	330
	55 - 64	11853	(a)	1842	424
	65+	9461	(a)	2964	254
Population by sex					
	Male	10478	(a)	6227	225
	Female	10307	(a)	6786	290

(a) Equivalised disposable income, being a variable derived on total disposable household income incorporates the sum of various components, independently of the pattern of gross/net collection and imputation/no imputation. It depends on a step using various components series, reflecting heterogeneous methods of imputation both in terms of algoritms and number of observations.

2.3 Non-sampling errors

2.3.1 Sampling frame and coverage errors

The new panel of the EU-SILC is a sub-sample of the Master Sample (MS) - the sampling frame used by the Statistics Portugal for household surveys.

The MS was designed and selected using the information of the last Census of Population and Housing (Census/2001). It is constituted by private dwellings and it excludes collective households and institutions since they represent 1% of the total population residing in Portugal.

The MS is constituted by almost 750 000 private dwellings (535 000 of which are as usual residence, the remaining are vacant, seasonal or for secondary use).

The MS is a stratified one-stage cluster sample. In each stratum the clusters were selected systematically with probability proportional to size (number of private dwellings of usual residence). The stratification was done at NUTS III level and the clusters are geographical areas constituted by one or more contiguous statistical sections (census enumeration areas).

Since the end of 2006 the MS is being updated. Each quarter a set of approximately 100 areas are updated in the field. There is no information about coverage problems.

2.3.2 Measurement and processing errors

2.3.2.1 Measurement errors

Different sources of measurement errors

The structure of the questionnaire was unchanged in 2009.

Same measurement errors persisted. These errors are basically associated with:

- The size of questionnaire, with a direct impact on an average duration of interview that exceeds an hour per household, producing mental fatigue and lack of attention during the annual interview and attrition on a year-to-year perspective.
- The complexity of income components collection, leading to misinterpretation and confusion between components – such as the one associated with old-age and survivors' benefits –, rough self-estimates by interviewed persons and missing or not credible values. In particular, distinguishing between gross and net income concepts is not easily perceived by interviewed persons and a special case of income – incomes that are not clearly classified in self-employment category or in employees' category – produces considerable longitudinal instability.
- Respondents were not receptive to the consultation of the annual tax income declaration.

Way the questionnaire was built up, field of testing, the effect of its design, content and wording

The structure of the questionnaire was unchanged in 2009.

Material deprivation questionnaire was built considering the respective regulation (regulation (EC) no. 362/2008 of 14 April 2008.

Definitions and recommendations from document EU-SILC 065 (2009 operation) were considered and whenever possible included as explanations throughout the questionnaire and fieldwork handbook.

Intensity and efficiency of interview training: number of training days, skills testing

Training was performed in two steps:

1st, fieldwork supervisors and regional technical managers had a one day training (4-5 May) by the core SILC team (concepts and consistence, software, collection rules);
2nd, supervisors and regional technical managers developed one day training (between 6th to 23th May, depending on the local office).

The majority of all new interviewers were followed by a supervisor, at least in one interview.

Information on studies, such as re-interviews, record check studies, or split-sample experiments

The supervision team controlled the quality of data collected, namely the number of missing values and unusual answers/situations, mainly by telephone contact (the exception being the personal control used in a specific region).

A thoroughly comparison with 2008 data was applied on income components and other variables such as age, sex, rotation and labour status. Also, and by income component, all outliers were examined. A comparative analysis with other sources and by income component was developed whenever available.

Results from models, for instance to assess the impact of using a financial year instead of a calendar year

No model was applied.

In particular there was no reporting on the use of a financial year different from the calendar year, which only occurs in a very few fiscal units related to international business groups and organised in accordance with corporate structures.

2.3.2.2 Processing errors

Data entry controls, coding controls, editing system applied to the data, main errors detected

Blaise is the software chosen to produce the CAPI application, which includes both questions and explanations and a package of prompt warnings and errors on the basis of ranges of feasible values and logical connections between questions. The original database gets attached a set of files of remarks by the interviewers in any unusual situation, making validation easier.

Coding experts, working in every household-addressed survey developed by Statistics Portugal, monitored the coding process.

Rates of failed edits for income variables

CAPI software includes several validation rules to prevent coherence errors, producing an immediate alert and correction during the interview.

A rate of failed edits is not available.

2.3.3 Non-response errors

2.3.3.1 Achieved sample size

Number of households for which an interview is accepted for the database

			Rotational group (DB075)									
	Total		2		3		4		1		Longitudinal component	
Number of valid addresses selected (DB120=11, 21, 22)	5707	100%	1023	100%	1276	100%	1330	100%	2078	100%	3629	100%
Interview accepted for database (DB135=1)	4961	92%	945	91%	1158	91%	1220	90%	1638	94%	3323	96%
Note: Longitudinal component correspond to rotational groups 2, 3 e 4.												

Number of persons of 16 years or older for which the interview is accepted for the database

	Rotational group (DB075)										Longitudinal component	
	Total		2		3		4		1			
Total	11163	100%	2202	100%	2604	100%	2726	100%	3631	100%	7532	100%
Information completed only from interview (RB250=11)	11101	99,4%	2198	99,8%	2588	99,4%	2705	99,2%	3610	99,4%	7491	99,5%
Individual unable to respond and no proxy possible (RB250=21)	30	0,3%	0	0,0%	5	0,2%	13	0,5%	12	0,3%	18	0,2%
Failed to return self-completed questionnaire (RB250=22)	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
Refusal to co-operate (RB250=23)	10	0,1%	1	0,0%	2	0,1%	2	0,1%	5	0,1%	5	0,1%
Person temporarely aw ay and no proxy possible (RB250=31)	18	0,2%	2	0,1%	7	0,3%	6	0,2%	3	0,1%	15	0,2%
No contact to other reasons (RB250=32)	4	0,0%	1	0,0%	2	0,1%	0	0,0%	1	0,0%	3	0,0%
Information not completed: reason unknow n (RB250=33)	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
Note: Longitudinal component correspond to rotational groups 2, 3 e 4.												

EU-SILC 2009 Longitudinal component												
	All individuals						Individuals aged 16+					
	Total		Sample persons (RB100=1)		Co-residents (RB100=2)		Total		Sample persons (RB100=1)		Co-residents (RB100=2)	
Total	8838	100%	7514	100%	1324	100%	7690	100%	7480	100%	210	100%
Not elegible person (RB250_F=-2)	1306	14,8%	186	2,5%	1120	84,6%	158	2,1%	152	2,0%	6	2,9%
Information completed only from interview (RB250=11)	7491	84,8%	7287	97,0%	204	15,4%	7491	97,4%	7287	97,4%	204	97,1%
Individual unable to respond and no proxy possible (RB250=21)	18	0,2%	18	0,2%	0	0,0%	18	0,2%	18	0,2%	0	0,0%
Failed to return self-completed questionnaire (RB250=22)	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
Refusal to co-operate (RB250=23)	5	0,1%	5	0,1%	0	0,0%	5	0,1%	5	0,1%	0	0,0%
Person temporarely aw ay and no proxy possible (RB250=31)	15	0,2%	15	0,2%	0	0,0%	15	0,2%	15	0,2%	0	0,0%
No contact to other reasons (RB250=32)	3	0,0%	3	0,0%	0	0,0%	3	0,0%	3	0,0%	0	0,0%
Information not completed: reason unknow n (RB250=33)	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%	0	0,0%
Note: including people w ho moved aw ay from household population (going abroad or to a colective household), i. e. RB110=5, and deceased people, i.e. RB110=6.												

Number of selected respondents (if applicable) for which the interview is accepted

Not applicable

2.3.3.2 Unit non-response

EU-SILC				
Response rate for households	2006	2007	2008	2009
Number of addresses successfully contacted (DB120=11)	4989	5243	4814	5641
Number of valid addressed selected (DB120=11, 21, 22)	5106	5380	4866	5707
Ra (address contact rate)	98%	97%	99%	99%
Number of household interviews completed and accepted for database (DB135=1)	4367	4310	4454	4961
Number of eligible households at contact addressed (DB130 filled)	4989	5243	4804	5641
Rh (proportion of complete households interviews accepted for database	88%	82%	93%	88%
NRh (household non-response rate)	14%	20%	8%	13%
Achieved sample size ratio	95%	99%	103%	111%
Number of household interviews completed and accepted for database (DB135=1) for longitudinal sample	3251	2986	3122	3323
Wave response rate	93%	92%	105%	106%
Longitudinal follow-up rate	87%	86%	96%	111%
Follow-up ratio	86%	96%		
Note: Indicators that are not applicable are indicated in grey.				

EU-SILC				
Response rate for persons	2006	2007	2008	2009
Number of personal interviews completed (RB250=11, 12, 13)	10148	9947	10101	11101
Number of eligible individuals in households whose interviews were completed ans accepted for the database (RB245=1, 2, 3)	10193	9988	10185	11163
Rp (proportion of complete personal interviews within the households accepted for the database)	100%	100%	99%	99%
Achieved sample size ratio	95%	98%	102%	110%
Number of personal interviews completed in t-1 and t (RB250=11, 12, 13) for sample persons	7388	6746	6938	7287
Number of personal interviews completed in t-1 and t (RB250=11, 12, 13) for co-residents	185	213	194	204
Number of personal interviews completed in t-1 and t (RB250=11, 12, 13)	7573	6959	7132	7491
Wave response rate	95%	92%	102%	105%
Sample persons selected excepted deceased	7571	6888	7156	7453
Longitudinal follow-up rate	98%	98%	97%	98%
Note: Indicators that are not applicable are indicated in grey.				

2.3.3.3 Distribution of households by 'record of contact at address', by 'household questionnaire result' and by 'household interview acceptance'

	Rotational group (DB075)							
Contact at address (DB120)	Total		2		3		4	
Total	3638	100%	1023	100%	1280	100%	1335	100%
Address contacted (DB120=11)	3598	99%	1017	99%	1259	98%	1322	99%
Address unable to access contacted (DB120=21)	31	1%	6	1%	17	1%	8	1%
Failed to return self-completed questionnaire (DB120=22)	0	0%	0	0%	0	0%	0	0%
Address does not exist or is a non-residential address or is unoccupied or not a principal residence (DB120=23)	9	0%	0	0%	4	0%	5	0%

	Rotational group (DB075)							
Household questionnaire result (DB130)	Total		2		3		4	
Total	3598	100%	1017	100%	1259	100%	1322	100%
Household questionnaire completed (DB130=11)	3323	92,4%	945	92,9%	1158	92,0%	1220	92,3%
Refusal to co-operate (DB130=21)	84	2,3%	15	1,5%	28	2,2%	41	3,1%
Entire household temporarily aw ay for duration of fieldw ork (DB130=22)	150	4,2%	48	4,7%	56	4,4%	46	3,5%
Household unable to respond (illness, inacapacity,…) (DB130=23)	33	0,9%	8	0,8%	14	1,1%	11	0,8%
Other reasons (DB130=24)	8	0,2%	1	0,1%	3	0,2%	4	0,3%

Household questionnaire result (DB110)	Rotational group (DB075)							
	Total		2		3		4	
Total	3709	100%	1045	100%	1301	100%	1363	100%
At the same address as last interview (DB110=1)	3508	95%	988	95%	1230	95%	1290	95%
Entire household moved to a private household within the country (DB110=2)	59	2%	20	2%	25	2%	14	1%
Entire household moved to a collective household or institution within the country (DB110=3)	15	0%	7	1%	4	0%	4	0%
Household moved outside the country (DB110=4)	11	0%	6	1%	0	0%	5	0%
Entire household died (DB110=5)	26	1%	7	1%	10	1%	9	1%
Household does not contain sample person (DB110=6)	7	0%	0	0%	4	0%	3	0%
Split-off household (DB110=8)	71	2%	15	1%	25	2%	31	2%
New adress added to the sample this wave or first wave (DB110=9)	0	0%	0	0%	0	0%	0	0%
Fusion (DB110=10)	1	0%	0	0%	1	0%	0	0%
Lost household (DB110=11)	11	0%	2	0%	2	0%	7	1%

For household interview acceptance (DB135), see item 2.3.3.1 (Achieved sample size).

2.3.3.4 Distribution of persons for membership status (RB110)

Household questionnaire result (RB110)	Rotational group (DB075)							
	Total		2		2		4	
Total	8841	100%	2555	100%	3104	100%	3182	100%
Was in this household in previous waves or current household member (RB110=1)	8463	96%	2463	96%	2959	95%	3041	96%
Moved into this household from another sample household since previous wave (RB110=2)	57	1%	13	1%	20	1%	24	1%
Moved into this household from outside sample since previous wave (RB110=3)	110	1%	30	1%	43	1%	37	1%
Newly born into this household since last wave (RB110=4)	44	0%	14	1%	16	1%	14	0%
Moved out since previous wave or last interview if not contacted in previous wave (RB110=5)	101	1%	24	1%	36	1%	41	1%
Died (RB110=6)	62	1%	11	0%	28	1%	23	1%
Lived in the household at least three months during the income reference period and was not recorded in the register of this household (RB110=7)	4	0%	0	0%	2	0%	2	0%

2.3.3.5 Item non-response

Item non-response is not available for Total disposable income (HY020), Total disposable income before social transfers other than old-age and survivors' benefits (HY022) and Total disposable income before all social transfers (HY023), because it corresponds to the sum of various components (the great majority of them corresponding themselves to the sum of various questions) independently of item non-response pattern.

Concerning this information component by component, information we attach counts of observations "Before imputation" and "After imputation".

2.4 Mode of data collection

EU-SILC 2009 longitudinal component												
	Rotational group (DB075)											
	Total		2		3		4		Sample persons (RB100=1)		Co-residents (RB100=2)	
Total	7491	100%	2198	100%	2588	100%	2705	100%	10897	100%	204	100%
Face to face interview : PAPI (RB260=1)	230	3%	88	4%	70	3%	72	3%	345	3%	7	3%
Face to face interview : CAPI (RB260=2)	5867	78%	1691	77%	2043	79%	2133	79%	8581	79%	124	61%
Proxy interview (RB260=5)	1394	19%	419	19%	475	18%	500	18%	1971	18%	73	36%

The distribution of household members aged 16 or over by data status (RB250) for each wave of EU-SILC longitudinal component was presented in the third table of item 2.3.3.1.

2.5 Imputation procedure

The net series of income data is obtained by the application of a specific gross-to-net micro simulation model[1]. This model was presented and is available on the Proceedings of the EU-SILC Conference, Helsinki, 6-8 November 2006, on Comparative EU Statistics on Income and Living Conditions: Issues and Challenges (Eurostat Methodologies and Working papers), pages 157-172, "Income in EU-SILC – Net/Gross Conversion Techniques for Building and Using EU-SILC Databases".

2.6 Imputed rent

In 2009, the imputed rent, i.e., the equivalent market rent to be paid for a similar dwelling, was calculated on the basis of a linear regression on HH070, dwelling dimension and degree of urbanization and with actual rents (HH060) as dependent variable.

2.7 Company cars

In the 2008 and 2009 questionnaires, this component was not collected individually, and integrated the variable on employees' non-monetary receipts.

[1] Carlos Farinha Rodrigues, PhD, ISEG/UTL and consultant of Statistics Portugal

3. COMPARABILITY

3.1 Basic concepts and definitions

There are no changes to be reported.

3.2 Components of income

3.2.1 Differences between the national definitions and standard EU-SILC definitions

No change to be reported in relation to the first wave Final Report.

3.2.2 The source or procedure used for the collection of income variables

No change to be reported in relation to the first wave Final Report.

3.2.3 The form in which income variables at component level have been obtained

The structure of the questionnaire in the income component part was deeply changed in 2008, with the aim of approaching the items requested to the annual tax income declaration. The structure of the questionnaire was unchanged in 2009.

3.2.4 The method used for obtaining income target variables in the required form (i.e. gross values)

It was collected according to doc. EU-SILC 065/09.

3.3 Tracing rules

Doc. 065/2009 rules were adopted.

4. *COHERENCE*

4.1 Comparison of income target variables and number of persons who receive income from each 'income component', with external sources

The objective of this section is to evaluate the results of the distribution of income, inequality and poverty obtained from the EU-SILC (2006, 2007, 2008, 2009) and from the HBS (2005).

When comparing the income structure of the EU-SILC and HBS, it is important to keep in mind the different concepts of income used in each survey. EU-SILC uses a monetary income concept, complemented with some categories of non-monetary income whereas the HBS uses the total Income concept, which includes both monetary and non-monetary income.

The differences on income structure will of course be reflected in the way income is distributed among individuals, as well as in different levels of inequality and poverty. The next table presents the indicators of inequality and poverty obtained by using each of the surveys. In the case of the HBS the first column (HBS1) is total income and the second one (HBS2) is monetary income. It is evident the impact of non-monetary income in the reduction of the risk-of-poverty rate, from 19% to 16%.

When comparing the income per adult equivalent distribution in 2005 estimated by EU-SILC 2006 and HBS 2005 outcomes are consistent, meaning for instance we get a poverty rate of 18.5% in EU-SILC and 19% in HBS.

When comparing income per adult equivalent in 2005 between EU-SILC 2006 and HBS 2005 we conclude that for the most relevant measures of income distribution, such as poverty rate and Gini coefficient, the HBS estimates are compatible with the EU-SILC confidence intervals.

					HBS1	HBS2
EU-SILC	**2006**	**2007**	**2008**	**2009**	**2005/2006**	
Income per adult equivalent	9.554 €	9.929 €	10.288 €	10.390 €	12.237 €	9.921 €
S80/S20	6.7	6.5	6.1	6.0	5,5	6,5
S90/S10	11.9	10.8	10.0	10.3	8,9	10,8
Gini index	37.7	36.8	35.8	35.4	34	37
Poverty line (60% of income per adult equivalent)	4.386 €	4.544 €	4.886 €	4.969 €	5.794 €	4.575 €
At-risk-of-poverty rate	18.5	18.1	18.5	17.9	16	19
Income reference year	**2005**	**2006**	**2007**	**2008**	**2005**	

Despite the inequality in the income distribution, the distance between the 20% of the population with the highest income (the top quintile) and the 20% of the population with the lowest income (the bottom quintile) have been gradually reduced from 6.7 in 2005 to 6.0 in 2008[2]. The evolution of the Gini coefficient in this period also reduced from 37.7 to 35.4, what confirms the tendency for the diminution of inequality in the income distribution.

Risk of poverty also reduced to 17.9% and inequality has continued the downward tendency.

EU-SILC 2009		
	PT	**EU 27**
Gini index	35.4	30.4
At-risk-of-poverty rate	17.9	16.3
Income reference year 2008		
Source of Data: Eurostat		
Date of extraction: 27 Dez 2011 13:31:47 CET		

In 2008, the risk-of-poverty rate in Portugal was 17.9%, more 1.6 p.p. than the mean for EU 27 (16.3%). The Gini index was 35.4%, more 5 p.p. than the Gini index for the EU 27 (30.4%).

[2] EU-SILC year n survey collects n-1 income data.